

# Adaptive Multilevel Monte Carlo Methods for Stochastic Variational Inequalities

Evgenia Babushkina<sup>1</sup> and Ralf Kornhuber<sup>1</sup>

<sup>1</sup>Freie Universität Berlin, Institut für Mathematik, Arnimallee 6, 14195 Berlin, Germany

## Abstract

While multilevel Monte Carlo (MLMC) methods for the numerical approximation of partial differential equations with uncertain coefficients enjoy great popularity, combinations with spatial adaptivity seem to be rare. We present an adaptive MLMC finite element approach based on deterministic adaptive mesh refinement for the arising "pathwise" problems and outline a convergence theory in terms of desired accuracy and required computational cost. Our theoretical and heuristic reasoning together with the efficiency of our new approach are confirmed by numerical experiments.

## 1 Introduction

Uncertainty quantification is a well-established and flourishing field in numerical analysis and scientific computing that connects theoretical challenges with a multitude of practical applications. While stochastic Galerkin approaches (cf., e.g., [4, 5, 45]) turned out as methods of choice for low dimensional uncertainties, Monte Carlo (MC) type of methods prove advantageous for high dimensional, highly nonlinear problems. While the classical MC method is very robust and extremely simple, sampling of stochastic data entails the numerical solution of numerous deterministic problems which makes performance the main weakness of this approach. A big step towards efficiency was made by Giles [26], who combined MC with multigrid techniques by introducing suitable hierarchies of subproblems associated with corresponding mesh hierarchies. Since then, multilevel Monte Carlo (MLMC) methods became a powerful tool in a variety of applications and its own field of mathematical research. We refer to elliptic problems with stochastic coefficients [9, 17, 18, 42], stochastic elliptic problems with multiple scales [1], parabolic stochastic problems [8], stochastic elliptic variational inequalities [37], and to [27] for a detailed overview.

Various approaches have been made to further enhance the efficiency of MLMC. For a given, quasi-uniform mesh hierarchy, Collier et al. [19] and Haji-Ali et al. [32] aim at reducing the computational work of MLMC by optimizing the actual selection of meshes from this hierarchy and other MLMC parameters.

Another, in a sense complementary approach to reduce the required computational work of MLMC is to apply adaptive mesh refinement techniques. Time discretization of an Itô stochastic differential equation by an self-adaptively chosen hierarchy of time steps has been suggested by Hoel et al. [33, 34] and a similar approach was presented by Gerstner and Heinz [25], including applications in computational finance.

Less appears to be known for partial differential equations with random coefficients. While a posteriori error estimation and adaptive mesh refinement have quite a history in finite

---

The authors want to thank Robert Scheichl for stimulating discussions. This work was supported by the German Ministry for Education and Research (BMBF).

element approximation of deterministic partial differential equations (cf., e.g., [2, 44]), related adaptive concepts for MLMC methods seem to be rare. Only recently, Eigel et al. [24] suggested an algorithm for constructing an adaptively refined hierarchy of meshes based on expectations of “pathwise” local error indicators and illustrated its properties by numerical experiments.

In this paper, we follow a novel approach, utilizing a whole family of different pathwise mesh hierarchies associated with different MC samples  $\omega \in \Omega$ . More precisely, for a given final tolerance  $Tol > 0$ , we choose a sequence of tolerances  $Tol_1 > \dots > Tol_L = Tol$  and approximate each of the different pathwise deterministic problems arising for each of the different samples  $\omega \in \Omega$  on each MLMC level  $l$  up to the accuracy  $Tol_l$  by finite elements on a different, adaptively refined “pathwise” mesh. We emphasize that any deterministic refinement strategy can be used for this purpose. The computation of sample averages is finally performed on an inductively constructed global mesh consisting of the union of simplices from all these pathwise meshes resulting from the different samples.

Utilizing well-known results on elliptic variational inequalities in a Hilbert space setting, we outline a general convergence theory in terms of the desired accuracy  $Tol$  and the required computational work. A framework for the efficient implementation of the resulting adaptive MLMC finite element method is provided by DUNE [12]. Numerical experiments illustrate our theoretical findings and the underlying heuristic reasoning. We also observe a significant reduction of computational work as compared to uniformly refined meshes.

The paper is organized as follows. After recalling some well-known existence and uniqueness results on elliptic variational inequalities, we first present a single-level MC method and its MLMC extension in a Hilbert space setting, together with general error estimates and upper bounds for the required computational work. These results rely on certain assumptions on the given pathwise hierarchy of approximating subspaces. Utilizing results from [37], we show in the next section that these assumptions are fulfilled for a uniform MLMC finite element method, i.e., in the case of uniform mesh refinement, and sufficiently smooth data. In Section 4 we present an adaptive MLMC method together with a convergence result exploiting related results for the deterministic case [15, 40]. Upper bounds for the computational work are discussed in light of existing deterministic optimality results for linear variational problems [11, 16, 38, 41]. Optimal bounds for the computational work are observed in the numerical experiments as reported in the final Section 5, but theoretical justification will be the subject of future research.

## 2 A stochastic variational problem

Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a complete probability space with  $\Omega$  denoting a sample space and let  $\mathcal{A} \in 2^\Omega$  be the  $\sigma$ -algebra of all possible events associated with a finite probability measure  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$  on  $\Omega$ . As usual,  $\mathbb{E}[\xi] = \int_\Omega \xi \, d\mathbb{P}$  describes the expectation of a random variable  $\xi$  and  $L^2(\Omega)$  denotes the Hilbert space of square integrable random variables on  $\Omega$ .

For a given separable Hilbert space  $H$ , equipped with the scalar product  $(\cdot, \cdot)_H$  and the associated norm  $\|\cdot\|_H = (\cdot, \cdot)_H^{1/2}$ , we introduce the Bochner-type space  $L^2(\Omega, \mathcal{A}, \mathbb{P}; H)$  of  $\mathbb{P}$ -measurable mappings  $v : \Omega \rightarrow H$  with the property  $\int_\Omega \|v\|_H^2 \, d\mathbb{P}(\omega) < \infty$ . We will use the abbreviation  $L^2(\Omega; H) = L^2(\Omega, \mathcal{A}, \mathbb{P}; H)$ . It is easily seen that  $L^2(\Omega; H)$  is also a Hilbert space with the scalar product

$$(v, w)_{L^2(\Omega; H)} = \int_\Omega (v, w)_H \, d\mathbb{P}(\omega), \quad v, w \in L^2(\Omega; H),$$

and the associated norm  $\|\cdot\|_{L^2(\Omega; H)} = (\cdot, \cdot)_{L^2(\Omega; H)}^{1/2}$ . The expectation in  $L^2(\Omega; H)$  is defined by

$$\mathbb{E}[v] = \int_\Omega v(\omega) \, d\mathbb{P}(\omega) \in H, \quad v \in L^2(\Omega; H).$$

Let  $a(\omega; \cdot, \cdot)$  and  $\ell(\omega; \cdot)$ ,  $\omega \in \Omega$ , denote families of bilinear forms and linear functionals on  $H$ , respectively. For a given subset  $K \subset H$  and any fixed realization  $\omega \in \Omega$ , we consider the “pathwise” variational inequality

$$u(\omega) \in K : \quad a(\omega; u(\omega), v - u(\omega)) \geq \ell(\omega; v - u(\omega)) \quad \forall v \in K. \quad (1)$$

Note that in the unconstrained case  $K = H$  the inequality (1) can be equivalently rewritten as the variational equality

$$u(\omega) \in H : \quad a(\omega; u(\omega), v) = \ell(\omega; v) \quad \forall v \in H. \quad (2)$$

**Assumption 2.1.** The subset  $K$  is non-empty, closed, and convex. For each realization  $\omega \in \Omega$  we have  $\ell(\omega; \cdot) \in H'$  and  $a(\omega; \cdot, \cdot)$  is bounded and coercive in the sense that

$$\gamma(\omega) \|v\|_H^2 \leq a(\omega; v, v), \quad a(\omega; v, w) \leq \Gamma(\omega) \|v\|_H \|w\|_H \quad \forall v, w \in H \quad (3)$$

holds with  $\gamma(\omega) \geq \gamma_0 > 0$  a.e. in  $\Omega$ , and  $\Gamma \in L^\infty(\Omega)$ . For all fixed  $v, w \in H$  the mappings  $a(\cdot; v, w)$ ,  $\ell(\cdot; v)$  are measurable and  $\ell \in L^2(\Omega; H')$ .

Assumption 2.1 yields existence, uniqueness, and regularity of pathwise solutions (cf., e.g., [35, Theorem 2.1] and [30, Proposition 1.2]).

**Proposition 2.1.** *Let Assumption 2.1 hold. Then the pathwise problem (1) admits a unique solution for each  $\omega \in \Omega$ , the solution map  $u : \Omega \mapsto H$  is measurable with respect to the Borel  $\sigma$ -algebra in  $H$ , and  $u \in L^2(\Omega; H)$ .*

Note that  $u \in L^2(\Omega; H)$  implies  $\mathbb{E}[u] \in H$ . It also follows that

$$u \in \mathcal{K} = \{v \in L^2(\Omega; H) \mid v(\omega) \in K \text{ a.e. in } \Omega\} \subset L^2(\Omega, H)$$

is the unique solution of the “mean-square” variational inequality

$$u \in \mathcal{K} : \quad \mathbb{E}[a(\cdot; u, v - u)] \geq \mathbb{E}[\ell(\cdot; v - u)] \quad \forall v \in \mathcal{K}. \quad (4)$$

To fix the ideas, we will often concentrate on the bilinear form

$$a(\omega; v, w) = \int_D \alpha(x, \omega) \nabla v \cdot \nabla w \, dx$$

and the functional

$$\ell(\omega; v) = \int_D f(x, \omega) \, dx$$

on the Sobolev space  $H = H_0^1(D)$  of weakly differentiable functions defined on a Lipschitz domain  $D \in \mathbb{R}^d$ ,  $d = 1, 2, 3$ , and the subset

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\}. \quad (5)$$

Note that random obstacles  $\chi \in L^2(\Omega, H_0^1(D))$  can be traced back to the case (5) by introducing the new variable  $w = u - \chi$ . For a detailed discussion of sufficient conditions on the coefficient  $\alpha$  and the right hand side  $f$  for existence and uniqueness of pathwise solutions, we refer to Section 4.

The remainder of this paper is devoted to the efficient approximation of the expected value  $\mathbb{E}[u]$  of the family of pathwise solutions  $u(\omega)$ ,  $\omega \in \Omega$ , of (1).

### 3 Pathwise Monte Carlo Galerkin methods

Our considerations will rely on basic properties of Monte Carlo methods as stated in the following lemma.

**Lemma 3.1.** *Let  $v \in L^2(\Omega; H)$ ,  $M \in \mathbb{N}$ , and let  $\omega_i \in \Omega$ ,  $i = 1, \dots, M$ , denote  $M$  independent, identically distributed samples. Then the Monte Carlo approximation*

$$\mathbb{E}_M[v] = \frac{1}{M} \sum_{i=1}^M v(\omega_i) \quad (6)$$

*of the expectation  $\mathbb{E}[v]$  satisfies the error estimate*

$$\|\mathbb{E}[v] - \mathbb{E}_M[v]\|_{L^2(\Omega; H)} \leq M^{-1/2} \|v\|_{L^2(\Omega; H)}.$$

*Proof.* As the samples  $\omega_i \in \Omega$  are independent and identically distributed, we have

$$\begin{aligned} \|\mathbb{E}[v] - \mathbb{E}_M[v]\|_{L^2(\Omega; H)}^2 &= \mathbb{E} \left[ \left\| \mathbb{E}[v] - \frac{1}{M} \sum_{i=1}^M v(\omega_i) \right\|_H^2 \right] = \frac{1}{M^2} \sum_{i=1}^M \mathbb{E} [\|\mathbb{E}[v] - v(\omega_i)\|_H^2] \\ &= \frac{1}{M} \mathbb{E} [\|\mathbb{E}[v] - v\|_H^2] = \frac{1}{M} (\mathbb{E}[\|v\|_H^2] - \|\mathbb{E}[v]\|_H^2) \leq \frac{1}{M} \|v\|_{L^2(\Omega; H)}^2. \end{aligned}$$

□

#### 3.1 Single-level methods

Our goal is to approximate the expectation  $\mathbb{E}[u]$  up to a prescribed tolerance  $Tol$ ,  $0 < Tol < 1$ . To this end, we select a subspace  $S(\omega)$  with finite dimension  $N(\omega)$  and a non-empty, closed, convex subset  $K_S(\omega) \subset S(\omega)$  for each  $\omega \in \Omega$ . We then consider the family of pathwise Galerkin approximations

$$u_S(\omega) \in K_S(\omega) : \quad a(\omega; u_S(\omega), v - u_S(\omega)) \geq \ell(\omega; v - u_S(\omega)) \quad \forall v \in K_S(\omega), \quad \omega \in \Omega. \quad (7)$$

**Assumption 3.1.** The set-valued map  $\Omega \ni \omega \mapsto K_S(\omega) \in H$  is measurable and there is a  $w \in L^2(\Omega, H)$  such that  $w(\omega) \in K_S(\omega)$  holds for all  $\omega \in \Omega$ .

In combination with Assumption 2.1, the Assumption 3.1 yields existence, uniqueness, and regularity of approximate pathwise solutions (cf., e.g., [31, Theorem 2.3 and 2.7]).

**Proposition 3.1.** *Let the Assumptions 2.1 and 3.1 hold. Then there is a unique solution  $u_S(\omega) \in K_S(\omega)$  of (7) for each  $\omega \in \Omega$ , the discretized solution map  $u_S(\omega) : \Omega \mapsto S(\omega) \subset H$  is measurable, and  $u_S \in L^2(\Omega; H)$ .*

Before we approximate the expectation  $\mathbb{E}[u]$  in terms of (approximations of)  $u_S(\omega)$ ,  $\omega \in \Omega$ , let us state some assumptions on  $u_S(\omega)$  and thus implicitly on the approximating family of spaces  $S(\omega)$ .

**Assumption 3.2.** The family  $u_S(\omega)$ ,  $\omega \in \Omega$ , satisfies the discretization error estimate

$$\|u - u_S\|_{L^2(\Omega; H)} \leq \frac{1}{4} Tol. \quad (8)$$

In general, the exact solution  $u_S(\omega)$  of variational inequality (7) is not available but can be only approximated up to a certain tolerance by an iterative solver.

**Assumption 3.3.** For each  $\omega \in \Omega$ , an approximate solution  $\tilde{u}_S(\omega) \in S(\omega)$  of the pathwise problem (7) can be computed with accuracy

$$\|u_S(\omega) - \tilde{u}_S(\omega)\|_H \leq \frac{1}{4} Tol, \quad \omega \in \Omega. \quad (9)$$

**Theorem 3.1.** *Let the Assumptions 2.1 and 3.1 - 3.3 hold. Then the inexact pathwise Monte Carlo Galerkin approximation*

$$\mathbb{E}_M[\tilde{u}_S] = \frac{1}{M} \sum_{i=1}^M \tilde{u}_S(\omega_i) \quad (10)$$

*of the expectation  $\mathbb{E}[u]$  satisfies the error estimate*

$$\|\mathbb{E}[u] - \mathbb{E}_M[\tilde{u}_S]\|_{L^2(\Omega;H)} \leq \frac{1}{2}Tol + M^{-1/2}(\|u\|_{L^2(\Omega;H)} + \frac{1}{2}Tol). \quad (11)$$

*In particular, the desired accuracy  $Tol$  is reached for*

$$M \geq C_0(u) Tol^{-2} \quad (12)$$

*with  $C_0(u) = (2\|u\|_{L^2(\Omega;H)} + 1)^2$ .*

*Proof.* The triangle inequality yields

$$\begin{aligned} & \|\mathbb{E}[u] - \mathbb{E}_M[\tilde{u}_S]\|_{L^2(\Omega;H)} \\ & \leq \|\mathbb{E}[u] - \mathbb{E}[u_S]\|_{L^2(\Omega;H)} + \|\mathbb{E}[u_S] - \mathbb{E}[\tilde{u}_S]\|_{L^2(\Omega;H)} + \|\mathbb{E}[\tilde{u}_S] - \mathbb{E}_M[\tilde{u}_S]\|_{L^2(\Omega;H)}. \end{aligned}$$

By the triangle inequality, the Cauchy-Schwarz inequality, and Assumption 3.2, we obtain

$$\|\mathbb{E}[u - u_S]\|_H \leq \mathbb{E}[\|u - u_S\|_H] \leq \|u - u_S\|_{L^2(\Omega;H)} \leq \frac{1}{4}Tol.$$

Similarly, Assumption 3.3 yields

$$\|\mathbb{E}[u_S - \tilde{u}_S]\|_H \leq \mathbb{E}[\|u_S - \tilde{u}_S\|_H] \leq \|u_S - \tilde{u}_S\|_{L^2(\Omega;H)} \leq \frac{1}{4}Tol.$$

From Lemma 3.1, the triangle inequality, and Assumptions 3.2 and 3.3, we finally get

$$\|\mathbb{E}[\tilde{u}_S] - \mathbb{E}_M[\tilde{u}_S]\|_{L^2(\Omega;H)} \leq M^{-1/2}\|\tilde{u}_S\|_{L^2(\Omega;H)} \leq M^{-1/2}(\|u\|_{L^2(\Omega;H)} + \frac{1}{2}Tol).$$

□

We now investigate the computational cost for the evaluation of  $E_M[\tilde{u}_S]$  with desired accuracy  $Tol$ . It is natural to measure the computational cost for the solution of the pathwise problem (7) in terms of the degrees of freedom  $N(\omega)$ .

**Assumption 3.4.** An approximation  $\tilde{u}_S(\omega)$  of the solution  $u_S(\omega)$  of (7) can be evaluated at computational cost bounded by

$$c_0(1 + \log(N(\omega)))^\mu N(\omega) \quad (13)$$

with positive constants  $c_0$  and  $\mu$  independent of  $Tol$ ,  $N(\omega)$ , and  $\omega \in \Omega$ .

In order to obtain a bound of the computational cost in terms of the desired accuracy,  $Tol$  has to be related to  $N(\omega)$ ,  $\omega \in \Omega$ .

**Assumption 3.5.** The dimension  $N(\omega)$  of the ansatz space  $S(\omega)$  providing the accuracy (8) satisfies

$$N(\omega) \leq c_1 Tol^{-s}, \quad \omega \in \Omega, \quad (14)$$

with positive constants  $c_1$  and  $s$  independent of  $N(\omega)$ ,  $Tol$ , and  $\omega \in \Omega$ .

Now we are ready to estimate the computational cost of the evaluation of  $\mathbb{E}(\tilde{u}_S)$  in terms of the desired accuracy  $Tol$ .

**Theorem 3.2.** *Let the Assumptions 2.1 and 3.1 - 3.5 hold. Then the inexact pathwise Monte Carlo Galerkin approximation  $\mathbb{E}_M(\tilde{u}_S)$  with accuracy  $Tol$  can be evaluated with computational cost bounded by*

$$C(1 - s \log(Tol))^\mu Tol^{-(s+2)} \quad (15)$$

with a positive constant  $C$  depending on  $c_0$ ,  $c_1$ ,  $\mu$ , and  $u$ .

*Proof.* Let  $M$  be the smallest positive integer such that (12) holds. Then, utilizing Theorem 3.1 and Assumption 3.4 the computational cost can be bounded according to

$$\begin{aligned} c_0 \sum_{i=1}^M (1 + \log(N(\omega)))^\mu N(\omega_i) &\leq c_0 c_1 M (1 + \log(c_1 Tol^{-s}))^\mu Tol^{-s} \\ &\leq C(1 + \log(Tol^{-s}))^\mu Tol^{-(s+2)} \end{aligned}$$

with a positive constant  $C$  depending on  $c_0$ ,  $c_1$ ,  $\mu$ , and  $u$ .  $\square$

### 3.2 Multilevel methods

For given initial tolerance  $0 < Tol_1 \leq 1$  and reduction factor  $q < 1$  we now define a sequence of tolerances by

$$Tol_l = q Tol_{l-1}, \quad l = 2, \dots, L, \quad (16)$$

with the final desired accuracy  $Tol = Tol_L$ . For each  $\omega \in \Omega$  we choose an associated hierarchy of subspaces  $S_l(\omega) \subset H$ , i.e.,

$$S_1(\omega) \subset S_2(\omega) \subset \dots \subset S_L(\omega) \subset H, \quad (17)$$

with finite dimensions  $N_l(\omega)$  and non-empty, closed, convex subsets  $K_l(\omega) \subset S_l(\omega)$ ,  $l = 1, \dots, L$ . We consider the family of pathwise Galerkin approximations

$$u_l(\omega) \in K_l(\omega) : \quad a(\omega; u_l(\omega), v - u_l(\omega)) \geq \ell(\omega; v - u_l(\omega)) \quad \forall v \in K_l(\omega), \quad \omega \in \Omega. \quad (18)$$

The following three assumptions are multilevel analogues of the Assumptions 3.1 - 3.3.

**Assumption 3.6.** For all  $l = 1, \dots, L$  the set-valued map  $\Omega \ni \omega \mapsto K_l(\omega) \in H$  is measurable and there is a  $w_l \in L^2(\Omega, H)$  such that  $w_l(\omega) \in K_l(\omega)$  holds for all  $\omega \in \Omega$ .

By Proposition 3.1, Assumptions 2.1 and 3.6 imply that there is a unique solution  $u_l(\omega)$  of (18) for each  $\omega \in \Omega$ , that the discretized solution map  $u_l(\omega) \mapsto S_l(\omega) \subset H$  is measurable, and that  $u_l \in L^2(\Omega; H)$  for all  $l = 1, \dots, L$ .

**Assumption 3.7.** For all  $l = 1, \dots, L$  the family  $u_l(\omega)$ ,  $\omega \in \Omega$ , satisfies the discretization error estimate

$$\|u - u_l\|_{L^2(\Omega; H)} \leq \frac{1}{4} Tol_l. \quad (19)$$

**Assumption 3.8.** For all  $l = 1, \dots, L$  and each  $\omega \in \Omega$ , an approximate solution  $\tilde{u}_l(\omega) \in S_l(\omega)$  of the pathwise problem (18) can be computed with accuracy

$$\|u_l(\omega) - \tilde{u}_l(\omega)\|_H \leq \frac{1}{4} Tol_l, \quad \omega \in \Omega. \quad (20)$$

We now prove an error bound for an inexact pathwise multilevel Monte Carlo Galerkin method.

**Theorem 3.3.** *Let the Assumptions 2.1 and 3.6 - 3.8 hold. Then the pathwise multilevel Monte Carlo Galerkin approximation*

$$\mathbb{E}^L[\tilde{u}_L] = \sum_{l=1}^L \mathbb{E}_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}] \quad (21)$$

of the expected value  $\mathbb{E}[u]$  with  $(M_l) \in \mathbb{N}^L$  and  $\tilde{u}_0 = 0$  satisfies the error estimate

$$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H)} \leq \frac{1}{2}Tol + M_1^{-1/2}\|u\|_{L^2(\Omega;H)} + \frac{1}{2}(1+q^{-1}) \sum_{l=1}^L M_l^{-1/2}Tol_l \quad (22)$$

In particular, the desired accuracy  $Tol$  is reached for the choice

$$M_1 \geq Tol^{-2} \max\{16 \|u\|_{L^2(\Omega;H)}^2, C_1 Tol_1^2\}, \quad M_l \geq C_1 l^{2(1+\varepsilon)} q^{-2(L-l)}, \quad l = 2, \dots, L, \quad (23)$$

with  $C_1 = 4(1+q^{-1})^2(1+\varepsilon^{-1})^2$  and some  $\varepsilon > 0$ .

*Proof.* Utilizing the triangle inequality together with Assumptions 3.7 and 3.8, we estimate the error according to

$$\begin{aligned} \|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H)} &\leq \|\mathbb{E}[u] - \mathbb{E}[u_L]\|_{L^2(\Omega;H)} + \|\mathbb{E}[u_L] - \mathbb{E}[\tilde{u}_L]\|_{L^2(\Omega;H)} + \|\mathbb{E}[\tilde{u}_L] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H)} \\ &\leq \frac{1}{2}Tol + \sum_{l=1}^L \|\mathbb{E}[\tilde{u}_l - \tilde{u}_{l-1}] - \mathbb{E}_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}]\|_{L^2(\Omega;H)}. \end{aligned}$$

Then, for  $l = 2, \dots, L$ , each term in the sum is bounded according to

$$\begin{aligned} \|\mathbb{E}[\tilde{u}_l - \tilde{u}_{l-1}] - \mathbb{E}_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}]\|_{L^2(\Omega;H)} &\leq M_l^{-1/2} \|\tilde{u}_l - \tilde{u}_{l-1}\|_{L^2(\Omega;H)} \\ &\leq M_l^{-1/2} (\|\tilde{u}_l - u_l\|_{L^2(\Omega;H)} + \|u_l - u\|_{L^2(\Omega;H)} + \|u - u_{l-1}\|_{L^2(\Omega;H)} + \|u_{l-1} - \tilde{u}_{l-1}\|_{L^2(\Omega;H)}) \\ &\leq \frac{1}{2}M_l^{-1/2}(Tol_l + Tol_{l-1}) = \frac{1}{2}(1+q^{-1})M_l^{-1/2}Tol_l \end{aligned}$$

by Lemma 3.1, Assumptions 3.7, 3.8 and (16). For  $l = 1$  and  $\tilde{u}_0 = 0$ , the same arguments yield

$$\|\mathbb{E}[\tilde{u}_1 - \tilde{u}_0] - \mathbb{E}_{M_1}[\tilde{u}_1 - \tilde{u}_0]\|_{L^2(\Omega;H)} \leq \frac{1}{2}M_1^{-1/2}Tol_1 + M_1^{-1/2}\|u\|_{L^2(\Omega;H)},$$

and the remaining assertion follows by elementary calculations utilizing  $Tol = q^{L-1}Tol_1$ .  $\square$

We now investigate the computational cost for the evaluation of  $\mathbb{E}^L[u_L]$ . The following assumptions are multilevel analogues of Assumptions 3.4 and 3.5.

**Assumption 3.9.** For all  $l = 1, \dots, L$  an approximation  $\tilde{u}_l(\omega)$  of the solution  $u_l(\omega)$  of (18) can be evaluated at computational cost bounded by

$$c_0(1 + \log(N_l(\omega)))^\mu N_l(\omega)$$

with  $c_0 > 0$ ,  $\mu > 0$  independent of  $Tol_l$ ,  $N_l(\omega)$ , and  $\omega \in \Omega$ .

**Assumption 3.10.** The dimension  $N_l(\omega)$  of the ansatz space  $S_l(\omega)$  providing the accuracy (19) satisfies

$$N_l(\omega) \leq c_1 Tol_l^{-s}, \quad \omega \in \Omega, \quad (24)$$

with positive constants  $c_1$ ,  $s$  independent of  $N_l(\omega)$ ,  $Tol_l$ , and  $\omega \in \Omega$ .

Now we are ready to state an upper bound for the computational cost for the evaluation of  $E^L[u_L]$  in terms of the desired accuracy  $Tol$ .

**Theorem 3.4.** *Let the Assumptions 2.1 and 3.6 - 3.10 hold. Then the inexact pathwise multilevel Monte Carlo Galerkin approximation  $\mathbb{E}^L(\tilde{u}_L)$  with accuracy  $Tol$  can be evaluated with computational cost bounded by*

$$C(1 - s \log Tol_1)^\mu Tol_1^{-\min\{2,s\}} Tol^{-\max\{2,s\}} \quad (25)$$

with a constant  $C$  depending on  $c_0, c_1, q, s, \mu, L, \varepsilon$ , and  $u$ .

*Proof.* Let  $M_l, l = 1, \dots, L$ , be the smallest positive integers such that (23) holds. Then, utilizing Theorem 3.3 together with Assumptions 3.9 and 3.10 the computational cost can be bounded according to

$$\begin{aligned} & c_0 \sum_{l=1}^L \sum_{i=1}^{M_l} ((1 + \log(N_l(\omega_i)))^\mu N_l(\omega_i) + (1 + \log(N_{l-1}(\omega_i)))^\mu N_{l-1}(\omega_i)) \\ & \leq c_0 c_1 \sum_{l=1}^L M_l ((1 + \log(c_1 Tol_l^{-s}))^\mu Tol_l^{-s} + (1 + \log(c_1 Tol_{l-1}^{-s}))^\mu Tol_{l-1}^{-s}) \\ & \leq C' \sum_{l=1}^L M_l (1 - s \log(Tol_l))^\mu Tol_l^{-s} \leq C'' (1 - s \log(Tol_1))^\mu \sum_{l=1}^L M_l Tol_l^{-s} \\ & \leq C (1 - s \log(Tol_1))^\mu Tol_1^{-2} Tol^{-s} \sum_{l=1}^L l^{2(1+\varepsilon)} q^{(s-2)(L-l)} \\ & = C (1 - s \log(Tol_1))^\mu Tol_1^{-s} Tol^{-2} \sum_{l=1}^L l^{2(1+\varepsilon)} q^{(2-s)(l-1)} \end{aligned}$$

with constants  $C', C''$ , and  $C$  depending on  $c_0, c_1, q, s, \mu, L, \varepsilon$ , and  $u$ .  $\square$

Observe that, roughly speaking, the inexact pathwise multilevel Monte Carlo Galerkin method is by a factor of  $Tol^{-\min\{s,2\}}$  faster than the single level version. Also note that Theorems 3.1 and 3.2 on pathwise single-level Monte Carlo Galerkin methods are contained as special cases of Theorems 3.3 and 3.4 for  $L = 1$ .

## 4 Multilevel Monte Carlo Finite Element methods

We consider problem (1) with the symmetric bilinear form

$$a(\omega; v, w) = \int_D \alpha(x, \omega) \nabla v(x) \cdot \nabla w(x) \, dx \quad (26)$$

and the linear functional

$$\ell(\omega; v) = \int_D f(x, \omega) v(x) \, dx, \quad (27)$$

both defined on the Sobolev space  $H = H_0^1(D)$  of weakly differentiable functions on a bounded Lipschitz domain  $D \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , equipped with the norm

$$\|v\|_H = \left( \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} v \right\|_{L^2(D)}^2 \right)^{1/2}.$$

The closed convex set  $K \in H$  of admissible solutions is given by

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\}. \quad (28)$$

We impose the following assumptions on the random coefficient  $\alpha$  and on the random right hand side  $f$ .



**Assumption 4.1.** The random diffusion coefficient  $\alpha$  and the right hand side  $f$  are strongly measurable mappings  $\Omega \ni \omega \mapsto \alpha(\cdot, \omega) \in L^\infty(D)$  and  $\Omega \ni \omega \mapsto f(\cdot, \omega) \in L^2(D)$  with the properties

$$0 < \alpha_- \leq \alpha(x, \omega) \leq \alpha_+ < \infty \quad \text{a.e. in } D \times \Omega, \quad (29)$$

and  $f \in L^2(\Omega, L^2(D))$ .

These assumptions imply Assumption 2.1 and thus existence and uniqueness of pathwise solutions  $u(\omega)$  of (1) and  $u \in L^2(\Omega, H)$ . Note that uniform coercivity (29) can be replaced by weaker conditions (cf., e.g., [37]).

Against the background of the general results from Section 3.2 we now concentrate on MLMC finite element methods, for the numerical approximation of the expectation  $\mathbb{E}[u]$ . Single level versions as derived in Section 3.1 are obtained for the special case  $L = 1$ .

#### 4.1 Uniform refinement

We assume for simplicity that  $D$  has a polygonal (polyhedral) boundary and consider the hierarchy of shape regular, conforming, quasiuniform partitions  $\mathcal{T}^{(k)}$ ,  $k \in \mathbb{N}$ , of  $D$  into simplices as obtained by subsequent uniform “red” refinement [7, 10, 13] of a given, intentionally coarse, initial partition  $\mathcal{T}^{(1)}$  (we will also assume that  $\mathcal{T}^{(1)}$  is sufficiently fine in a sense to be specified below). Then

$$h_k = \max_{t \in \mathcal{T}^{(k)}} \text{diam}(t) = 2^{-k} h_1, \quad k \in \mathbb{N},$$

and the associated finite element spaces

$$S^{(k)} = \{v \in H \mid v|_t \text{ is affine } \forall t \in \mathcal{T}^{(k)}\}, \quad k \in \mathbb{N}, \quad (30)$$

form a hierarchy of subspaces of  $H$ . We consider the pathwise approximations  $u^{(k)}(\omega) \in K^{(k)} = S^{(k)} \cap K$  characterized by

$$a(\omega; u^{(k)}(\omega), v - u^{(k)}(\omega)) \geq \ell(\omega; v - u^{(k)}) \quad \forall v \in K^{(k)}, \quad \omega \in \Omega. \quad (31)$$

**Assumption 4.2.** The spatial domain  $D$  is convex and the random coefficient  $\alpha$  is a measurable map  $\Omega \ni \omega \mapsto \alpha(\cdot, \omega) \in C^1(\bar{D})$  with the property  $\alpha \in L^\infty(\Omega; C^1(\bar{D}))$ .

The following discretization error estimate is taken from [37, Proposition 4.3].

**Theorem 4.1.** *Let the Assumptions 4.1 and 4.2 hold. Then the error estimate*

$$\|u - u^{(k)}\|_{L^2(\Omega; H^1(D))} \leq C(a, f) h_k \quad (32)$$

*holds with a positive constant  $C(a, f)$  independent of  $\omega \in \Omega$  and  $k \in \mathbb{N}$ .*

We make sure that  $\mathcal{T}^{(1)}$  is fine enough to guarantee

$$\|u - u^{(1)}\|_{L^2(\Omega; H^1(D))} \leq \frac{1}{4} \text{Tot}_1 \quad (33)$$

by selecting  $h_1$  such that  $C(a, f) h_1 \leq \frac{1}{4} \text{Tot}_1$  and define a uniform MLMC hierarchy in the sense of (17) according to

$$S_l(\omega) = S^{(r(l-1)+1)}, \quad K_l = S_l(\omega) \cap K, \quad l = 1, \dots, L, \quad \omega \in \Omega. \quad (34)$$

Then Assumption 3.6 is trivially satisfied and Theorem 4.1 implies the accuracy Assumption 3.7 by choosing  $r \in \mathbb{N}$  such that  $2^{-r} \leq q$ . Furthermore, Assumption 3.8 can be satisfied by sufficiently many steps of any iterative solver for elliptic variational inequalities that converges uniformly in  $\omega$  and consists of basic arithmetic or max operations, thus

preserving measurability (cf., e.g., [20, 29, 36, 39, 43]). Then, by Theorem 3.3, the resulting uniform, inexact MLMC finite element approximation  $\mathbb{E}^L[\tilde{u}_L]$  with the number  $M_l$  of MC samples on each level chosen according to (23) satisfies the desired error estimate

$$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H)} \leq Tol. \quad (35)$$

It is well-known (cf. [37, Section 4.5], [6, Corollary 4.1]) that Standard Monotone Multigrid (STDMMG) methods [36, 39] satisfy Assumption 3.9 with  $\mu = 4$  in  $d = 1$  space dimension and with  $\mu = 5$  in  $d = 2$  space dimensions. In spite of computational evidence, no theoretical justification of mesh-independent convergence rates seem to be available for  $d = 3$ . Finally, utilizing again Theorem 4.1, we find that Assumption 3.10 holds with  $s = d$ , because the dimension  $N_l$  of  $S_l$  is bounded by  $h_{r(l-1)+1}^{-d}$  and thus by  $Tol_l^{-d}$  up to a constant. Hence, Theorem 3.4 implies the following result on the efficiency of uniform MLMC finite element methods.

**Theorem 4.2.** *Let the Assumptions 4.1, 4.2, and (33) hold, let  $M_l$ ,  $l = 1, \dots, L$ , be the smallest positive integers such that (23) holds and let STDMMG be used for the iterative solution of the pathwise discretized obstacle problems of the form (18). Then the resulting uniform MLMC finite element method provides an approximation  $\mathbb{E}^L(\tilde{u}_L)$  with prescribed accuracy (35) at computational cost bounded by*

$$C(1 - d \log Tol_1)^\mu Tol_1^{-\min\{2, d\}} Tol^{-\max\{2, d\}}$$

with a constant  $C$  independent of  $Tol_1$ ,  $\mu = 4$  in  $d = 1$  space dimension, and  $\mu = 5$  in  $d = 2$  space dimensions.

## 4.2 Adaptive refinement

Let  $\mathcal{T}_{uni}^{(k)}$  be the partition of  $D$  with vertices  $\mathcal{N}_{uni}^{(k)}$  as obtained by  $k - 1 = 0, 1, \dots$  uniform refinements of the initial partition  $\mathcal{T}_{uni}^{(1)} = \mathcal{T}^{(1)}$  and let  $S_{uni}^{(k)}$  denote the corresponding piecewise affine finite element space spanned by the nodal basis functions  $\lambda_p^{(k)}$ ,  $p \in \mathcal{N}_{uni}^{(k)}$ . We collect all nodal basis functions occurring on all uniform refinement levels in the set

$$\Lambda = \{\lambda_p^{(j)} \mid p \in \mathcal{N}_{uni}^{(j)}, j \in \mathbb{N}\}.$$

For each fixed  $\omega \in \Omega$ , we now use a deterministic adaptive strategy to construct a nested sequence of finite element spaces

$$S^{(1)}(\omega) \subset \dots \subset S^{(k)}(\omega) \quad \text{with} \quad S^{(k)}(\omega) = \text{span } \Lambda^{(k)}(\omega), \quad k \in \mathbb{N}, \quad (36)$$

by successive, adaptive selection of a hierarchy of finite subsets

$$\Lambda^{(1)}(\omega) \subset \dots \subset \Lambda^{(k)}(\omega) \subset \Lambda, \quad k \in \mathbb{N}, \quad \text{with} \quad \Lambda^{(1)}(\omega) = \{\lambda_p^{(1)} \mid p \in \mathcal{N}_{uni}^{(1)}\}.$$

Note that each finite subset  $\Lambda^{(k)}(\omega)$  can be associated to a partition  $\mathcal{T}^{(k)}(\omega)$  of  $D$  with so-called “hanging nodes” [28, Section 3.1]. We assume measurability and pointwise convergence in the following sense.

**Assumption 4.3.** The adaptively refined finite element spaces  $S^{(k)}(\omega)$ ,  $k \in \mathbb{N}$  are measurable and  $\|u(\omega) - u^{(k)}(\omega)\|_H \rightarrow 0$  for  $k \rightarrow \infty$  holds for each fixed  $\omega \in \Omega$ .

Existing adaptive algorithms as suggested, e.g., by Siebert and Veeser [40], Braess et al. [15, Section 5] (see also Carstensen [16]) provide convergence for each fixed  $\omega \in \Omega$ . We only sketch that also measurability and thus Assumption 4.3 is satisfied by these algorithms, referring to [3] for details.

For given measurable  $S^{(k)}(\omega) = \text{span } \Lambda^{(k)}(\omega)$  the approximate solution  $\tilde{u}^{(k)}(\omega) \in S^{(k)}(\omega)$  of the discretized pathwise problem (31) is measurable and gives rise to local error indicators

$\eta_t(\omega)$ ,  $t \in \mathcal{T}^{(k)}(\omega)$ . Usual local error indicators  $\eta_t(\omega)$  are obtained by basic algebraic operations of the residual and thus inherit measurability. Those simplices that sufficiently contribute to the overall error estimate

$$\eta^{(k)} = \left( \sum_{t \in \mathcal{T}^{(k)}(\omega)} \eta_t^2 \right)^{1/2} \quad (37)$$

are marked for refinement. Usual marking strategies (cf., e.g., Dörfler [23]) preserve measurability, thus providing a measurable extended subset  $\Lambda^{(k+1)}(\omega)$  and a corresponding measurable space  $S^{(k+1)} = \text{span } \Lambda^{(k+1)}(\omega)$ .

Assuming in addition that  $\mathcal{T}_{uni}^{(1)} = \mathcal{T}^{(1)}$  is fine enough to guarantee (33), we define an adaptive MLMC hierarchy in the sense of (17) according to

$$S_1(\omega) = S_{uni}^{(1)}, \quad S_l(\omega) = S^{(k_l(\omega))}(\omega), \quad l = 2, \dots, L, \quad (38)$$

where  $k_l(\omega)$  is the smallest natural number such that

$$\|u(\omega) - u^{(k_l(\omega))}(\omega)\|_H \leq \frac{1}{4} Tol_l. \quad (39)$$

and  $Tol_l$  is chosen according to (16). Note that  $k_l(\omega)$  might not be uniformly bounded in  $\omega \in \Omega$ .

In order to show Assumption 3.6, i.e., measurability of  $S_l(\omega)$ , we first observe that  $S^{(k+1)}(\omega) = S^{(k)}(\omega) + V^{(k)}(\omega)$  with measurable incremental spaces

$$V^{(k)}(\omega) = \text{span } \Lambda^{(k+1)} \setminus \Lambda^{(k)}, \quad k \in \mathbb{N}.$$

Measurability of  $S^{(k)}(\omega)$  also implies measurability of  $u^{(k)}(\omega)$  and thus measurability of

$$S_l(\omega) = S^{k_l(\omega)} = S^{(1)} + \bigcup_{k \in \mathbb{N}} H(\|u(\omega) - u^{(k)}(\omega)\|_H - \frac{1}{4} Tol_l) V^{(k)}(\omega),$$

where  $H(\cdot)$  stands for Heaviside step function.

Assumption 3.7 holds by construction and Assumption 3.8 can be satisfied by sufficiently many steps of any iterative solver for elliptic variational inequalities that converges uniformly in  $\omega$  and consists of basic arithmetic or max operations, thus preserving measurability (cf., e.g., [20, 29, 36, 39, 43]). Hence, the following convergence result is a consequence of Theorem 3.3.

**Theorem 4.3.** *Let the Assumption 4.1, 4.3, and (33) hold, and let  $M_l$ ,  $l = 1, \dots, L$ , be the smallest positive integers such that (23) is satisfied. Then the adaptive MLMC finite element method based on the MLMC hierarchy defined in (38) provides an approximation  $\mathbb{E}(\tilde{u}_L)$  with accuracy*

$$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H)} \leq Tol. \quad (40)$$

Let us briefly discuss the required computational cost (cf. Theorems 3.4 and 4.2). For fixed  $\omega \in \Omega$  and unconstrained problems, i.e.,  $K = H$ , the optimality condition

$$N_l(\omega) \leq c_1(\omega) Tol_l^{-d} \quad (41)$$

has been established for a variety of adaptive refinement strategies (cf. e.g., [11, 41, 38]). Uniform upper bounds for  $c_1(\omega)$  as required in Assumption 3.10 and corresponding upper bounds for the computational cost as stated in Theorem 3.4 are observed in the numerical experiments to be presented in the next section. Theoretical validation will be the subject of future research.

## 5 Numerical Experiments

In this section we investigate the adaptive MLMC finite element approach presented in the preceding sections from a numerical perspective.

Starting from an initial tolerance  $Tol_1 > 0$  to be specified later, we choose  $Tol_l$  according to (16) with  $q = \frac{1}{2}$  and  $Tol = Tol_L$  for various  $L \in \mathbb{N}$ . Similar to the approach in [37, Remark 4.11], we compute the number of samples  $M_l$ ,  $l = 1, \dots, L$ , by rounding up the solution of the optimization problem

$$(M_1, \dots, M_L) = \arg \min_{\underline{M} \in \mathbb{N}^L} \sum_{l=1}^L Tol_l^{-d} M_l, \quad (42)$$

subject to the constraints

$$\sum_{l=1}^L M_l^{-\frac{1}{2}} Tol_l \leq \frac{1}{6} Tol, \quad M_1^{-\frac{1}{2}} \|u\|_{L^2(\Omega; H)} \leq \frac{1}{4} Tol, \quad M_l \geq 1 \quad \forall l = 1, \dots, L. \quad (43)$$

Note that the exact solution  $u$  and thus  $\|u\|_{L^2(\Omega; H)}$  is known in our numerical experiments. More sophisticated selections, e.g., along the lines of [19, 32] seem to be promising but are beyond the scope of this paper.

Deterministic adaptive refinement is performed as suggested by Siebert and Veiser [40] involving Dörfler marking [23] with  $\theta = 0.6$ . The error indicators  $\eta_t(\omega)$  are given by local contributions to the hierarchical error estimator of the form (37) as suggested in [46, Theorem 3.5]. Here, the exact finite element solution is replaced by an approximation provided by an iterative method to be described below. In the unconstrained case  $K = H$ , this approach is reducing to the classical hierarchical error estimation (cf., e.g., [14, 21] or [22, Section 6.1.4]). Note that the error is estimated in the energy norm. We use local “red” mesh refinement [7, 10, 13] with hanging nodes [28, Section 3.1] as implemented in the finite element toolbox DUNE [12].

The initial accuracy condition (33) is addressed by setting

$$Tol_1 = \frac{4}{\sigma_{disc} \alpha_-} \mathbb{E}_{1000}[\eta^{(1)}]$$

with a safety factor  $\sigma_{disc} = 0.5$  accounting for error estimation and inexact evaluation of the discrete solution. The accuracy criterion (39) is replaced by the approximation

$$\eta^{(k_l(\omega))}(\omega) \leq \frac{1}{4} \sigma_{disc} \alpha_- Tol_l.$$

Discretized variational inequalities of the form (31) are solved iteratively by truncated non-smooth Newton multigrid methods (TNNMG) [28, 29] with nested iteration, because TNNMG is easier to implement and usually converges faster than STDMMG [28]. Numerical experiments (see, e.g., [37, Section 5.]) also indicate that TNNMG satisfies Assumption 3.9 with  $\mu = 0$ . Note that both STDMMG and TRCMG reduce to classical multigrid with Gauß-Seidel smoothing in the unconstrained case  $K = H$ . The accuracy condition (20) is replaced by the uniform stopping criterion

$$\|u_{\nu+1}^{(k)} - u_{\nu}^{(k)}\|_H \leq \frac{1}{4} \sigma_{alg} Tol$$

with  $u_{\nu}^{(k)}$  denoting the  $\nu$ -th iterate, a safety factor  $\sigma_{alg} = 0.1$  accounting for estimating the algebraic error  $\|u^{(k)} - u_{\nu}^{(k)}\|_H$  by  $\|u_{\nu+1}^{(k)} - u_{\nu}^{(k)}\|_H$ , and the final tolerance  $Tol = Tol_L$ . In view of the above mentioned optimal convergence properties of TNNMG, the cost for each algebraic solution in  $S_l(\omega)$  is set to the corresponding number of unknowns  $N_l(\omega) = \dim S_l(\omega)$ . Hence, the computational cost for the adaptive MLMC method with  $L$  levels is given by

$$cost_L = \sum_{l=1}^L \sum_{i=1}^{M_l} N_l(\omega_i)$$

which reduces to  $cost_L = \sum_{l=1}^L N_l M_l$  in case of uniform refinement.

Table 1: Number of samples  $M_1, \dots, M_L$ ,  $L = 1, \dots, 6$ , for Poisson problem,  $\beta = 10$ .

$L$	1	2	3	4	5	6
$M_1$	573	2290	9161	36864	230400	1327104
$M_2$		64	1026	9217	57600	331776
$M_3$			256	2304	14400	82944
$M_4$				577	3600	20736
$M_5$					900	5184
$M_6$						1296

### 5.1 Poisson equation with uncertain right-hand side

We consider the Poisson problem

$$u(\omega) \in \{w \in H^1(D) \mid w|_{\partial D} = g(\omega)\} : \quad a(\omega; u(\omega), v) = \ell(\omega; w) \quad \forall v \in H_0^1(D) \quad (44)$$

with  $D = (-1, 1)^2$  in  $d = 2$  space dimensions, the bilinear form

$$a(\omega; v, w) = \int_D \nabla v \cdot \nabla w \, dx, \quad v, w \in H, \quad (45)$$

the right hand side

$$\ell(\omega; v) = \int_D f(x, \omega) v \, dx, \quad v \in H, \quad (46)$$

with uncertain source term

$$f(x, \omega) = e^{-\beta|x-Y(\omega)|^2} (4\beta^2|x-Y(\omega)|^2 - 4\beta), \quad (47)$$

and uncertain, inhomogeneous boundary conditions

$$g(x, \omega) = e^{-\beta|x-Y(\omega)|^2} \quad x \in \partial D.$$

Here,  $\beta$  is a positive constant and  $Y(\omega) = (Y_1(\omega), Y_2(\omega))^T$  is a stochastic vector whose components are uniformly distributed random variables  $Y_1, Y_2 \sim \mathcal{U}(-0.25, 0.25)$ . For each  $\omega \in \Omega$  a pathwise solution of (44) is given by

$$u(x, \omega) = e^{-\beta|x-Y(\omega)|^2}, \quad x \in D. \quad (48)$$

As Assumption 4.1 is satisfied, this solution is unique and we have spatial regularity in the sense that  $u \in L^2(\Omega, H^2(D))$  (cf. Assumption 4.2). However,  $u(\omega)$  exhibits a peak at  $(Y_1(\omega), Y_2(\omega)) \in D$  that becomes more pronounced with increasing  $\beta$  thus leading to larger constants  $C(a, f)$  in the uniform error estimate (32).

We will compare the performance of multilevel Monte Carlo finite element methods as presented in the preceding Section 4 based on uniform and adaptive refinement for  $\beta = 10, 50, 150$ . The initial partition  $\mathcal{T}^{(1)}$  is obtained by applying three uniform refinement steps to the partition of the unit square  $\overline{D}$  into two congruent triangles with right angles at  $(1, -1)$  and  $(-1, 1)$ . The numbers of MLMC samples  $M_l$ ,  $l = 1, \dots, L$ , for different values of  $L$  as computed from constrained minimization (42) for different values of  $\beta$  are reported in Tables 1 - 3. We emphasize that these values are used both for uniform and adaptive MLMC.

Figure 1 illustrates the convergence properties of uniform and adaptive MLMC methods for different values of  $\beta$ . Here, the statistical error  $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H^1(D))}$  is approximated by a Monte Carlo method utilizing 10 independent realisations  $\|E[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$ . In all cases adaptive MLMC method is slightly more accurate than the uniform version. For all values  $\beta = 10, 50, 150$  both uniform and adaptive MLMC match the desired accuracy  $Tol_L$  as indicated by the dotted line, thus nicely confirming our theoretical results

Table 2: Number of samples  $M_1, \dots, M_L$ ,  $L = 1, \dots, 6$ , for Poisson problem,  $\beta = 50$ .

$L$	1	2	3	4	5	6
$M_1$	230	918	5184	36864	230401	1327105
$M_2$		99	1297	9211	57600	331776
$M_3$			325	2304	14400	82944
$M_4$				577	3600	20737
$M_5$					901	5184
$M_6$						1297

Table 3: Number of samples  $M_1, \dots, M_L$ ,  $L = 1, \dots, 6$ , for Poisson problem,  $\beta = 150$ .

$L$	1	2	3	4	5	6
$M_1$	128	576	5184	36864	230400	1327104
$M_2$		144	1296	9216	57599	331776
$M_3$			324	2304	14399	82944
$M_4$				576	3600	20736
$M_5$					900	5184
$M_6$						1296

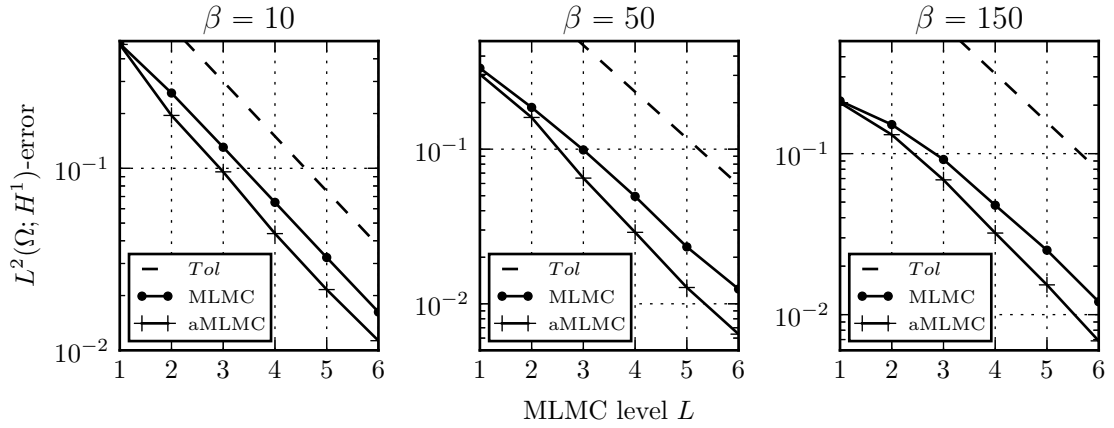


Figure 1: Statistical error of uniform and adaptive MLMC over the number of levels  $L$  for the Poisson problem.

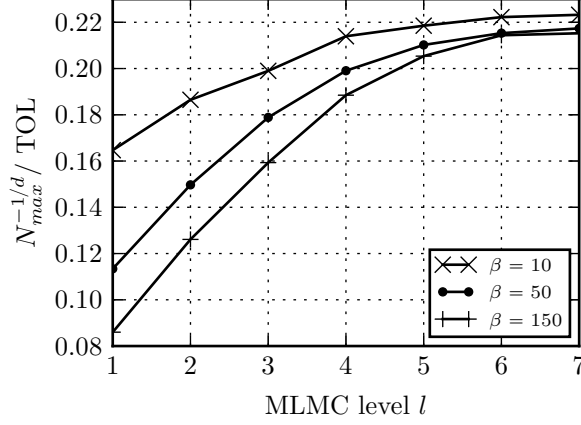


Figure 2: Ratio  $N_{l,\max}^{-1/d}/Tol_l$ ,  $d = 2$ , over the MLMC levels  $l = 1, \dots, 7$  for the Poisson problem.

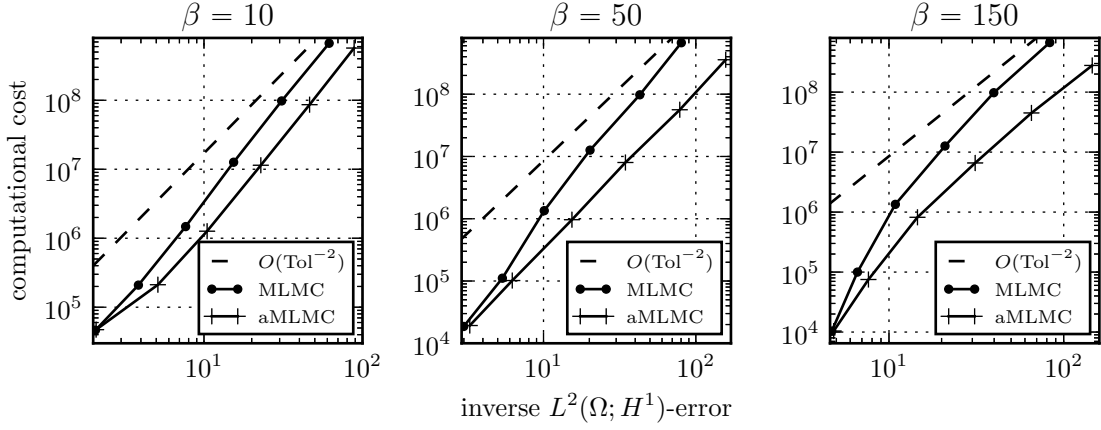


Figure 3: Computational cost of uniform and adaptive MLMC over the inverse  $L^2(\Omega; H^1)$ -error for the Poisson problem.

(cf. Theorem 4.1 and 4.3) also in this slightly more general case. Note that for  $\beta = 150$ , i.e. for quite pronounced peaks, uniform MLMC needs more levels to reach the asymptotic regime than adaptive MLMC.

Upper bounds of the computational cost of MLMC in terms of the desired accuracy  $Tol_L$  as stated in Theorem 3.4 strongly rely on Assumption 3.10 postulating  $N_l(\omega) = \mathcal{O}(Tol_l^{-s})$ . While, under suitable regularity conditions, Assumption 3.10 holds with  $s = d$  for uniform MLMC there is no theoretical evidence for adaptive MLMC. In order to check Assumption 3.10 for adaptive MLMC numerically, Figure 2 displays the ratio  $N_{l,\max}^{-1/d}/Tol_l$ ,  $d = 2$ , with  $N_{l,\max} = \max_{i=1,\dots,I} N_l(\omega_i)$  for  $I = 1000$  independent samples  $\omega_i$  over the levels  $l = 1, \dots, 7$ . For all three values of  $\beta$  this ratio seems to saturate at about 0.22, indicating that Assumption 3.10 with  $s = d = 2$  also holds for adaptive MLMC in this case.

On this background, we expect from Theorem 3.4 that the computational cost both of uniform and adaptive MLMC asymptotically behaves like  $\mathcal{O}(Tol_L^{-2})$ . Figure 3 shows the computational cost over the inverse of the corresponding  $L^2(\Omega; H^1)$ -error together with the expected asymptotic behavior (dotted line). While the asymptotic regime is clearly reached by adaptive MLMC for  $\beta = 150$ , 50, and almost reached or  $\beta = 10$ , this is not the case for the uniform version. Observe also the gain of efficiency by adaptive mesh refinement which is increasing with increasing  $\beta$ .

Table 4 finally reports on the average mesh sizes on the final level  $L = 1, \dots, 6$  in uniform

Table 4: Average number of unknowns on MLMC levels  $L$  for the Poisson problem.

$L$	1	2	3	4	5	6
uniform	81	289	1089	4225	16641	66049
adapted $\beta = 10$	85	285	833	3433	13013	44619
adapted $\beta = 50$	84	167	466	1534	5345	20469
adapted $\beta = 150$	82	119	333	986	3223	12132

and adaptive MLMC. While for  $\beta = 10$  the corresponding uniform and adaptive mesh sizes stay relatively close to each other, corresponding meshes for uniform MLMC are by a factor of about 5 larger than for adaptive MLMC. Even though most of the work in MLMC methods is performed on coarser levels, this again reflects the gain of efficiency by adaptive mesh refinement.

## 5.2 Obstacle problem with uncertain diffusion coefficient and right-hand side

We consider an elliptic variational inequality of the form (1) with  $D = (0, 1)$  in  $d = 1$  space dimension,

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\} \subset H, \quad H = H_0^1(D),$$

the bilinear form

$$a(\omega; v, w) = \int_D \alpha(x, \omega) \nabla v \cdot \nabla w \, dx, \quad v, w \in H, \quad (49)$$

with uncertain diffusion coefficient

$$\alpha(x, \omega) = 1 + \frac{\cos x^2}{10} Y_1(\omega) + \frac{\sin x^2}{10} Y_2(\omega), \quad (50)$$

and the right hand side

$$\ell(\omega; v) = \int_D f(x, \omega) \, dx, \quad v \in H, \quad (51)$$

with uncertain source term

$$f(x, \omega) = \begin{cases} -8e^{2(Y_1(\omega)+Y_2(\omega))} (a(x, \omega) \cdot (3x^2 - r^2) \\ \quad + (x^2 - r^2)x^2 \left( -\frac{\sin x^2}{10} Y_1(\omega) + \frac{\cos x^2}{10} Y_2(\omega) \right)) , & x > r \\ 4r^2 e^{2(Y_1(\omega)+Y_2(\omega))} (a(x, \omega) \cdot (-1 - r^2 + x^2) \\ \quad + (-2 - 2r^2 + x^2)x^2 \left( -\frac{\sin x^2}{10} Y_1(\omega) + \frac{\cos x^2}{10} Y_2(\omega) \right)) , & x \leq r \end{cases}$$

denoting

$$r = r(Y_1(\omega), Y_2(\omega)) = 0.7 + \frac{Y_1(\omega) + Y_2(\omega)}{10}.$$

Here,  $Y_1, Y_2 \sim \mathcal{U}(-1, 1)$  stand for uniformly distributed random variables. For each  $\omega \in \Omega$  a solution of the corresponding pathwise problem (1) is given by

$$u(x, \omega) = \max \{ (x^2 - r^2) e^{Y_1(\omega)+Y_2(\omega)}, 0 \}^2, \quad x \in D.$$

As Assumption 4.1 is satisfied, this solution is unique and we have  $u \in L^2(\Omega, H)$ .

We will compare the numerical behavior of MLMC finite element methods with uniform and adaptive spatial mesh refinement as presented in Section 4. The initial partition  $\mathcal{T}^{(1)}$  of  $\overline{D} = [0, 1]$  consists of eight closed intervals with length  $1/8$ . The numbers of MLMC



Table 5: Number of samples  $M_1, \dots, M_L$ ,  $L = 1, \dots, 8$ , for the obstacle problem.

$L$	1	2	3	4	5	6	7	8
$M_1$	39	464	3384	19694	101614	487224	2226149	9837306
$M_2$		184	1343	7816	40325	193355	883448	3903937
$M_3$			533	3102	16003	76732	350596	1549278
$M_4$				1231	6350	30451	139134	614831
$M_5$					2520	12084	55215	243996
$M_6$						4795	21912	96829
$M_7$							8695	38426
$M_8$								15249

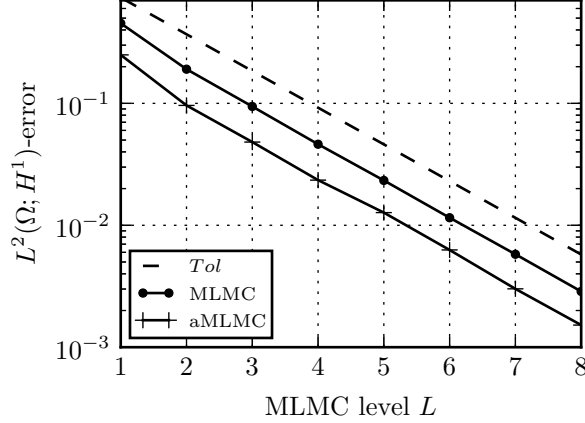


Figure 4: Statistical error of uniform and adaptive MLMC versus the number of levels  $L$  for the obstacle problem.

samples  $M_l$ ,  $l = 1, \dots, L$ , for  $L = 1, \dots, 8$  as computed from constrained minimization (42) are reported in Table 5. These values are used both for uniform and adaptive MLMC. Figure 4 shows the evolution of the statistical error  $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H^1(D))}$  of uniform and adaptive MLMC over the final level  $L = 1, \dots, 8$ . As in the previous numerical experiment, the exact error  $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H^1(D))}$  is approximated by a Monte Carlo method utilizing 10 independent realisations  $\|E[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$ . Adaptive MLMC appears to be slightly more accurate than the uniform version. As expected from Theorem 4.1 and 4.3, both for uniform and adaptive MLMC the error is bounded by the prescribed tolerance  $Tol_L$  indicated by the dotted line.

As the given data clearly satisfy Assumption 4.2, the general Assumption 3.10 holds true for uniform MLMC. Hence, Theorem 4.2 provides the upper bound  $\mathcal{O}(Tol_L^{-2})$  for the computational cost of uniform MLMC. As corresponding theoretical evidence is still missing for adaptive MLMC, we check Assumption 3.10 numerically. Figure 5 displays the ratio  $N_{l, \max}^{-1/d}/Tol_l$  with  $d = 1$  and  $N_{l, \max} = \max_{i=1, \dots, I} N_l(\omega_i)$  for  $I = 1000$  independent samples  $\omega_i$  over the levels  $l = 1, \dots, 10$ . This ratio seems to saturate at about 0.18, indicating that Assumption 3.10 with  $s = d = 1$  also holds for adaptive MLMC in this case.

From Theorem 4.2 and Theorem 3.4, combined with numerical evidence as depicted in Figure 5, we expect that the computational cost both of uniform and adaptive MLMC asymptotically behaves like  $\mathcal{O}(Tol_L^{-2})$ . This is confirmed by Figure 3 showing the computational cost over the inverse of the corresponding  $L^2(\Omega; H^1)$ -error together with the expected asymptotic behavior (dotted line). We also observe a considerable gain of efficiency of adaptive MLMC as compared to the uniform version which is also reflected by the average size of the meshes on the final levels as reported in Table 6.

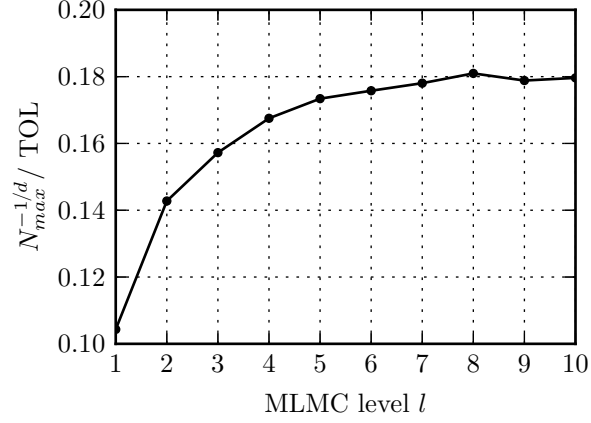


Figure 5: Ratio  $N_{l,\max}^{-1/d}/\text{Tol}_l$ ,  $d = 1$ , over the MLMC levels  $l = 1, \dots, 10$  for the obstacle problem.

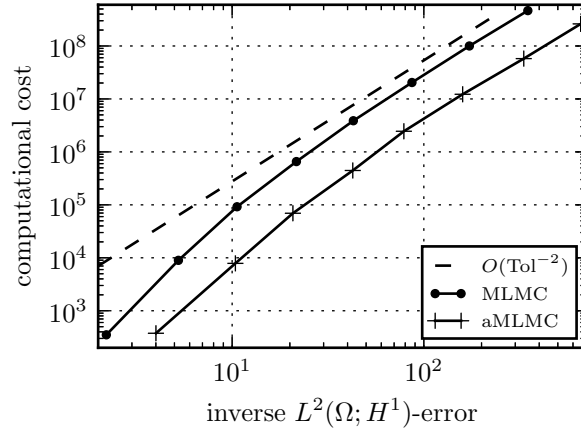


Figure 6: Computational cost of uniform and adaptive MLMC the inverse  $L^2(\Omega; H^1)$ -error for the obstacle problem.

Table 6: Average number of unknowns on MLMC levels  $L$  for the obstacle problem.

$L$	1	2	3	4	5	6	7	8
uniform	9	17	33	65	129	257	513	1025
adapted	10	12	16	24	40	72	134	260

## References

- [1] A. Abdulle, A. Barth, and C. Schwab. Multilevel Monte Carlo methods for stochastic elliptic multiscale pdes. *Multiscale Modeling and Simulation*, 11(4):1033–1070, 2013.
- [2] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Wiley, 2011.
- [3] E. Babushkina. *Adaptive Finite Element Methods for Elliptic Variational Problems with Uncertain Coefficients*. PhD thesis, FU Berlin, in preparation.
- [4] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Rev.*, 52:317–355, 2010.
- [5] I. Babuska, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42:800–825, 2004.
- [6] L. Badea. Global convergence rate of a standard multigrid method for variational inequalities. *IMA J. Numer. Anal.*, pages 1–20, 2013.
- [7] R.E. Bank, A.H. Sherman, and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In R. Stepleman, editor, *Scientific Computing*, pages 3–17. IMACS/North Holland, 1983.
- [8] A. Barth, A. Lang, and C. Schwab. Multilevel Monte Carlo method for parabolic stochastic partial differential equations. *BIT Numerical Mathematics*, 53(1):3–27, 2013.
- [9] A. Barth, C. Schwab, and N. Zollinger. Multi-level monte carlo finite element method for elliptic pdes with stochastic coefficients. *Numerische Mathematik*, 119(1):123–161, 2011.
- [10] J. Bey. Simplicial grid refinement: on Freudenthal’s algorithm and the optimal number of congruence classes. *Numer. Math.*, 85(1):1–29, 2000.
- [11] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97:219–268, 2004.
- [12] M. Blatt, A. Burchardt, A. Dedner, Ch. Engwer, J. Fahlke, B. Flemisch, Ch. Gersbacher, C. Gräser, F. Gruber, Ch. Grüninger, D. Kempf, R. Klöfkorn, T. Malkmus, S. Müthing, M. Nolte, M. Piatkowski, and O. Sander. The distributed and unified numerics environment, version 2.4. *Archive of Numerical Software*, 4(100):13–29, 2016.
- [13] F. Bornemann, B. Erdmann, and R. Kornhuber. Adaptive multilevel methods in three space dimensions. *Int. J. Numer. Meth. Engrg.*, 36:3187–3203, 1993.
- [14] F. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM Journal on Numerical Analysis*, 33:3:1188–1204, 1996.
- [15] D. Braess, C. Carstensen, and R.H.W. Hoppe. Convergence analysis of a conforming adaptive finite element method for an obstacle problem. *Numer. Math.*, 107:455–471, 2007.
- [16] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Methods Appl. Math.*, 67:1195–1253, 2014.

- [17] J. Charrier, R. Scheichl, and A. Teckentrup. Finite element error analysis of elliptic pdes with random coefficients and its application to multilevel Monte Carlo methods. *SIAM Journal on Numerical Analysis*, 51(1):322–352, 2013.
- [18] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic pdes with random coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.
- [19] N. Collier, A.-L. Haji-Ali, F. Nobile, E. von Schwerin, and R. Tempone. A continuation multilevel Monte Carlo algorithm. *BIT Numerical Mathematics*, 55:399–432, 2015.
- [20] R. W. Cottle and J.-S. Pang. *The Linear Complementarity Problem*. SIAM, 1992.
- [21] P. Deuffhard, P. Leinen, and H. Yserentant. Concepts of an adaptive hierarchical finite element code. *IMPACT Comput. Sci. Eng.*, 1(1):3–35, March 1989.
- [22] P. Deuffhard and M. Weiser. *Adaptive Numerical Solution of PDEs*, volume 55. de Gruyter, Berlin, 2012.
- [23] W. Dörfler. A convergent adaptive algorithm for poisson’s equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.
- [24] M. Eigel, Ch. Merdon, and J. Neumann. An adaptive multi level Monte-Carlo method with stochastic bounds for quantities of interest with uncertain data. *WIAS Preprint 2060*, 2015.
- [25] T. Gerstner and S. Heinz. Dimension- and time-adaptive multilevel Monte Carlo methods. In J. Garcke and M. Griebel, editors, *Sparse Grids and Applications, Lecture Notes in Computational Science and Engineering, Volume 88*, pages 107–120. Springer, 2013.
- [26] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.
- [27] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259, 2015.
- [28] C. Gräser. *Convex minimization and phase field models*. PhD thesis, FU Berlin, 2011.
- [29] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comput. Math.*, 27:1–44, 2009.
- [30] J. Gwinner. A class of random variational inequalities and simple random unilateral boundary value problems—existence, discretization, finite element approximation. *Stochastic Anal. Appl.*, 18(6):967–993, 2000.
- [31] J. Gwinner and F. Raciti. On a class of random variational inequalities on random sets. *Numer. Func. Anal. Opt.*, 27:619–636, 2006.
- [32] A.-L. Haji-Ali, F. Nobile, E. v. Schwerin, and R. Tempone. Optimization of mesh hierarchies in multilevel Monte Carlo samplers. *Stoch. PDE: Anal. Comp.*, 4:76–112, 2016.
- [33] H. Hoel, E. v. Schwerin, A. Szepessy, and R. Tempone. Implementation and analysis of an adaptive multilevel Monte Carlo algorithm. *Monte Carlo Methods and Applications*, 20:1–41, 2014.
- [34] H. Hoel, E. Von Schwerin, A. Szepessy, and R. Tempone. Adaptive multilevel Monte Carlo simulation. In *Numerical Analysis of Multiscale Computations*, pages 217–234. Springer, 2012.

- [35] D. Kinderlehrer and G. Stampaccia. *Introduction to variational inequalities and their applications*. Academic Press, New York, 1980.
- [36] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math.*, 69:167–184, 1994.
- [37] R. Kornhuber, C. Schwab, and M.-W. Wolf. Multi-level Monte-Carlo finite element methods for stochastic elliptic variational inequalities. *SIAM Journal on Numerical Analysis*, 52(3):1243–1268, 2014.
- [38] C. Kreuzer and K. G. Siebert. Decay rates of adaptive finite elements with dörfler marking. *Numerische Mathematik*, 117:679–716, 2011.
- [39] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Opt.*, 11:77–95, 1984.
- [40] K. Siebert and A. Veiser. A unilaterally constrained quadratic minimization with adaptive finite elements. *SIAM J. Optim.*, 18:260–289, 2007.
- [41] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comp. Math.*, 7:245–269, 2007.
- [42] A. Teckentrup, R. Scheichl, M. Giles, and E. Ullmann. Further analysis of multi-level Monte Carlo methods for elliptic pdes with random coefficients. *Numerische Mathematik*, 125(3):569–600, 2013.
- [43] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. SIAM, 1992.
- [44] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley, 1996.
- [45] D. Xiu and G. E. Karniadakis. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002.
- [46] Q. Zou, A. Veiser, R. Kornhuber, and C. Gräser. Hierarchical error estimates for the energy functional in obstacle problems. *Numer. Math.*, 117(4):653–677, 2011.